DOI:
10.1159/000458770

Document Version
Peer reviewed version

Measuring fatigue using the Multidimensional Fatigue Inventory (MFI-20): a questionable factor structure in haemodialysis patients

Joseph Chilcot[1], Ayman Guirguis[2,3], Karin Friedli[3], Michael Almond[4], Andrew Davenport[5], Clara Day[6], David Wellsted[3], Ken Farrington[2,7]

[1]Institute of Psychiatry, Psychology and Neuroscience, King's College London; [2]Postgraduate Medical School, University of Hertfordshire; [3]Centre for Lifespan and Chronic Illness Research, University of Hertfordshire; [4]Southend University Hospital NHS Foundation Trust; [5]UCL Centre for Nephrology, Royal Free Hospital NHS Foundation Trust; [6]Department of Renal Medicine, University Hospitals Birmingham NHS Foundation Trust; [7]Lister Hospital, East & North Hertfordshire NHS Trust

Corresponding author: Dr Joseph Chilcot (joseph.chilcot@kcl.ac.uk). Health Psychology Section, Institute of Psychiatry, Psychology and Neuroscience King's College London, 5th floor Bermondsey Wing, Guy's Hospital Campus, London Bridge, London, SE1 9RT

Key words: fatigue; confirmatory factor analysis; haemodialysis; kidney disease; Multidimensional Fatigue Inventory; MFI-20

Running Head: Factor structure of the MFI-20 in HD patients

Funding: This work was supported by the National Institute for Health Research programme, Research for Patient Benefit programme (PB-PG-0110-21073)

Conflict of interest: The authors declare no conflicts of interest

ABSTRACT

Background/aims: Fatigue is recognised as a common and burdensome symptom among dialysis patients. A growing body of research is devoted to understanding fatigue in advanced kidney disease, yet its measurement is challenging within this context. Our aim was to evaluate the *factor structure* underlying the Multidimensional Fatigue Inventory (MFI-20) and to examine associations with clinical factors and mood.

Methods: Data was evaluated for confirmatory factor analysis (CFA) from the screening phase of a multicentre randomised placebo-controlled trial of sertraline in HD patients. 470 patients completed the MFI-20, which purports to measure five components of fatigue (general, mental and physical fatigue, reduced motivation and reduced activity). CFA models were evaluated in MPlus 7.3 using Robust Maximum Likelihood (MLR) estimation.

Results: Evaluation of the original five factors revealed low internal reliability for the general factor and reduced activity, and high intercorrelations between all sum scores. CFA revealed poor model fit for the original 5-factor MFI-20 model (CFI=0.738; TLI= 0.689; RMSEA= 0.101). Alternative models, including 1, 3 and bi-factor models all demonstrated poor fit to the data. No reliable factor model was confirmed prohibiting the examination of factors associated with fatigue.

Conclusions: We were not able to confirm the factor structure of the MFI-20 in a large sample of HD patients. Certain items may lack suitable face validity in this context.

INTRODUCTION:

Fatigue is a burdensome and common symptom in patients with End-Stage Kidney Failure (ESKF) which is associated with poor outcomes [1], including survival [2]. Accordingly, there is growing recognition regarding the importance of understanding fatigue in ESKF [3-5] and developing evidence based interventions to ameliorate symptoms. It is important however, that the measurement of fatigue is thoroughly evaluated to ensure the reliability and validity of study findings [1,6].

Estimates of fatigue in ESKF vary across studies partly due to different self-report fatigue severity being used [1]. In psychometric analysis, a technique call Confirmatory Factor Analysis (CFA) is commonly used to examine the underlying structure of a questionnaire, that is, to evaluate the number of purported factors a questionnaire has. CFAs are known as measurement models in that they describe how observed variables (called indicators or items) measure underlying directly unobservable constructs (called latent factors or just factors). CFA is employed to test how well the latent factors explain observed variables (i.e. how well the model fits the data). CFA differs from exploratory factor analysis (EFA) because in CFA the researcher has an idea of the proposed measurement structure and seeks to confirm that structure. Accordingly, CFA allows researchers to evaluate the structural validity of a measure.

Fatigue has been described as a multifactorial problem, yet many multidimensional self report measures are in fact sufficiently unidimensional to warrant use of a total score (i.e. the sum of all of items to form one severity score) as a reliable measure of general fatigue severity [7]. For example, we recently showed in both renal [8] and multiple sclerosis [6] populations that the Chalder Fatigue

Questionnaire is suitably unidimensional to warrant a total severity score, rather than consisting of separate subfactors for "mental" and "physical" fatigue.

A growing number of studies, including renal samples [9,10], have used the Multidimensional Fatigue Inventory (MFI-20) [11] to evaluate five proposed components (i.e. factors) of fatigue, namely: *general fatigue, mental and physical fatigue, reduced motivation and reduced activity*. In this measure, each of the factors are measured by four items which are summed, with higher scores on each representing great severity of that construct. Studies in other long-term conditions [11-14] and general population samples [15] have supported the factor structure of the MFI-20 [11], suggesting it measures these five related constructs (or components) of fatigue. However, no study to date has examined the factor structure of the MFI-20, using CFA, in a sample of ESKF on haemodialysis (HD). Our study objective was to test the MFI-20's factor structure (i.e. does it measure five factors of fatigue?) in a large UK sample of HD patients, and if supported, examine the associations (correlates) with depression symptoms, demographic and clinical factors.

METHODS:

Design: The present study utilises screening data to select patients into a multicentre placebo controlled feasibility randomised control trial (RCT) of sertraline in HD patients with mild to moderate Major Depressive Disorder. The full RCT protocol (trial registration number: ISRCTN06146268) [16], and the outcome paper [17], have been published elsewhere. Screening occurred in 709 HD patients of which a subsample (n=470) completed the MFI-20. Ethical approval was granted from the National Health Service ethics committee.

Patients: The study recruited HD patient across five UK dialysis centers. All patients that had been receiving HD treatment for a minimum of 3 months and could speak and read English well enough to complete the questionnaires were eligible to be screened. For full details of the original RCT, see [16].

Fatigue measure: Fatigue was assessed using the Multidimensional Fatigue Inventory (MFI-20) [11]. The MFI-20 contains a total of 20 items, purporting to measure five factors; *general* (items 1, 5, 12, 16), *mental* (items 7, 11, 13, 19) and *physical fatigue* (items 2, 8, 14, 20), *reduced motivation* (items 4, 9, 15, 18) and *reduced activity* (items 3, 6, 10, 17). Each factor contains 4-items and is scored on a 5 point Likert type scale, where participants rate each item according to their agreement ("yes, that is true" to disagreement "no, that is not true"). Items scores are summed to create a sum score for each of the five factors. Items 2, 5, 9, 10, 13, 14, 16-19 were recoded so that higher scores represent more fatigue.

Depression Measures: Depression symptoms were measured using two validated screening tools, 1) Beck Depression Inventory-II (BDI-II) [18] and 2) Patient Health Questionnaire-9 (PHQ-9) [19]. These measures have been shown to perform well in HD patients [20-22].

Clinical and demographic factors: Clinical data was collected from medical records, which included the comorbidities (presence of diabetes, heart disease, stroke, cancer, limb amputation, liver disease, lung disease), dialysis vintage (length of time on dialysis; months), haemoglobin (g/L), serum albumin (g/L) and dialysis treatment

adequacy (Kt/V). C-reactive protein (CRP, mg/L) was available in 213 patients. Demographic information was collected through a self-report questionnaire.

Statistical methods: The factor structure of the MFI-20 was examined using CFA in MPlus 7.3. Competing factor models for the MFI-20 (i.e. different factor solutions) were estimated using Robust Maximum Likelihood Estimation (MLR) since item responses were slightly skewed. Missing data was minimal with 1% or less data missing for each item, corresponding to 0.09% of missing data across the entire MFI-20 dataset.

The original 5-factor model was tested along with a 1-factor models (with all 20 items loaded onto one general factor) and a bi-factor model. In the bi-factor model, all 20 items were loaded onto a general fatigue factor. In addition, respective items were also loaded five group factors (i.e. the original 5-factors), with correlations between each of these latent factors fixed to zero, and variances of the latent factors fixed to 1.

In CFA, a non-significant chi-square is desired which suggests that the reproduced and observed model covariance matrixes do not differ significantly meaning that the data fits the proposed model structure. However the chi-square statistic is sensitive to sample size [23] and therefore should not be used alone to determine the appropriateness of model fit. Accordingly, assessment of goodness-of-fit based on standard criteria was also examined. The following fit statistics and their standard cut-offs (used to indicate adequate fit) were evaluated: the root mean squared error of approximation (RMSEA) <.08, confirmatory fit index (CFI) >.95, and Tucker-Lewis index (TLI) >.95 [24].

RESULTS:

Patient sample characteristics:

The mean age of the sample was 63.8 (16.6) years (median=67; interquartile range=25), with most patients white (n=292; 62.1%) and male (n=307; 65.3%). The median time on dialysis was 22 months (interquartile range= 42; min=3; max=495). Comorbidity and clinical data was representative of a dialysis population. Heart disease and diabetes was present in 140 (29.8%) and 166 (35.3%) of patients respectively. 45 patients (9.6%) had cancer, 14 (3%) liver disease and 30 (6.4%) lung disease.

Mean dialysis adequacy (Kt/V) was 1.5 (0.3). Mean haemoglobin and serum albumin levels were 11.7 (1.3) g/L and 37.4 (4.7) g/L respectively. Median CRP was 5.0 (interquartile range= 13) mg/L.

Mean depression scores on the BDI and PHQ-9 were 13.4 (s.d=11.1) and 6.9 (s.d=6.2) respectively.

Summary of the MFI-20 sum scores:

Figure 1 displays the MFI-20 individual item means (95% confidence intervals). Means, intercorrelations and internal reliabilities for the original five MFI-20 sum scores are shown in table 1. Both the general and reduced activity had poor internal reliabilities, since Cronbach's alpha were some way below 0.7. Intercorrelations revealed a strong association between the general fatigue factor with both reduced activity and reduced motivation ($rs > 0.90$, $ps<0.01$), casting doubt about the uniqueness of this factors as distinct constructs since they share a significantly large amount of variance. Furthermore, both reduced activity and

motivation correlated very highly (r=0.97, p<0.01). The MFI-20 total score (sum score of all 20-items) did have good internal reliability (α=.89).

Confirmatory factor analysis (CFA):

CFA revealed that the original 5-factor model had poor model fit (model a, table 2), as evidenced by inadequate fit criteria. In order to try and specify which items/factors could be contributing to this poor fit, five separate CFAs for each individual factor were evaluated independently. None of the individual factors analyzed separately demonstrated suitable fit as indicated by the fit indices (see table 2) and all had significant chi-square statistics (ps> 0.05).

An alternative 1-factor model, with all items loaded onto one general factor indicating a total score (model b, table 2) had poor fit, as did a bi-factor model (model c). Given these poor model fits, and the high correlations between the general fatigue factor with both reduced activity and reduced motivation, an additional 3-factor model was tested where both reduced activity and motivation items were loaded onto the general fatigue factor. This 3-factor model also revealed poor model fit (model d, table 2). Due to the unreliability of the factor structure, it was not appropriate to examine factor correlations (convergent validity) with depression, demographic and clinical factors since no valid latent fatigue dimensions were determined from the CFA.

Sensitivity of the estimator (MLR)

Given the poor model fit, all CFA models were rerun using an alternative estimator; Weighted Least-Squares with mean and variance adjustment (WLSMV). WLSMV produces unbiased, consistent and efficient parameter estimates and standard errors where ordinal responses are used [25]. All the factor models described

in table 2 produced similar results when using the WLSMV estimator. For example, the original five factor model (CFI=0.86; TLI=0.82; RMSEA=0.12) and bi-factor model (CFI=0.81; TLI=0.78; RMSEA=0.14) displayed poor model fit. Fit was also poor for all the alternative CFA models tested.

DISCUSSION

The present study aimed to evaluate the factor structure of the MFI-20, a commonly used self-report fatigue measure which was originally designed to measure five constructs or components of fatigue. There is considerable recognition regarding the importance of understanding the fatigue within ESKF, given its adverse effects upon outcomes [1]. However, the most appropriate tools to measure self reported fatigue within dialysis patients remains unknown [5]. Understanding which tool has the most robust psychometric properties is important if we are to select the most suitable tool in studies that measure fatigue as either an outcome or as a predictor of outcomes.

We found that the correlations between the original five sum scores were high, which casts doubt about the uniqueness of this factors as distinct constructs since they share a significantly large amount of variance (table 1). Furthermore, two factors (general and reduced activity) had poor internal reliability, suggesting a lack of consistency (low item correlations) between the items of these factors.

Confirmatory factor models testing the original five-factor MFI-20 model structure proved to have poor model fit, suggesting that the purported MFI-20 factor is not supported in HD patients. Alternative CFA models including a bi-factor model and a 3 factor model also had poor model fit (table 2). Therefore, our analysis failed

to support a reliable factor structure of the MFI-20 as applied to HD patients, suggesting that this measure has low validity in this setting.

The lack of a reliable factor structure maybe the result of comprehension difficulties in HD patients [9] , high item and factor correlations and potentially low face validity for some items. Although no formal evaluation of face validity was conducted as part of this study, several items appear potentially problematic as indicators of fatigue. For example *"I think I do very little in a day"* (item 10), *"I feel fit"* (item 1) and *"I am rested"* (item 12) are likely to relate to the high morbidity observed in the dialysis population, rather than fatigue per se. Item 20 *"Physically I feel I am in an excellent condition"* also appears to be related more to a general reflection about health status. Accordingly, detailed evaluation of the face validity of the MFI-20 in dialysis patients should be encouraged, including both patient and health care professional input. As indicated elsewhere, tailoring MFI-20 items responses could improve comprehension [9] and ultimately improve the content validity.

Other reasons for the poor model fit observed here may be because the MFI-20 attempts to measure multiple components of fatigue (severity, consequences and antecedent factors), which is likely complicated in dialysis patients due to the high prevalence of depressive symptoms [26], co-morbidities, inflammation and varied post dialysis recovery between individuals.

Our results conflict with past studies supporting the factor structure of MFI-20 across a range of settings [11-13]. Some of these studies however relied upon relatively small samples sizes and exploratory factor analytic methods [13,14], which may account for discrepancies between study findings. Others have failed to confirm the original factor structure in cancer patients, suggesting instead a modified 3-factor

model, albeit with marginal fit (CFI=0.882; RMSEA=0.079). Given the difference between patient samples and analytic methods used, a meta-confirmatory factor analysis of the MFI-20 may be useful to further evaluate the measures structural validity. Given the importance of understanding fatigue in ESKF further efforts to establish which current measures provide the best validity and reliability are needed, and potentially a renal specific fatigue tool developed.

Study limitations

As discussed, the lack of a structured evaluation of face validity is a limitation of our study. We did not plan to perform a full psychometric analysis, so re-test reliability (i.e. correlational performance over time) and convergent validity (correlational strengthen with another fatigue measure) were not assessed. We therefore encourage a full psychometric evaluation of the MFI-20 in kidney patients, in order to evaluate its measurement properties further. However, establishing a reliable factor structure is often a prerequisite for additional psychometric analysis.

With regards to our sample it was representative in terms of age (median=67) and comorbidity profile to the UK HD population [27,28]. However, dialysis vintage was relatively low (median=22 months), although we do not expect this to impact upon our findings since this is unlikely to lead to model invariance and there is mixed findings regarding the association between dialysis vintage and fatigue severity [1].

Conclusions

To conclude, we were not able to confirm the factor structure of the MFI-20 in a large sample of HD patients. It is possible that certain items lack suitable face validity in this context. Currently, the *factor structure* of the MFI-20 in dialysis patients remains questionable.

REFERENCES

1 Artom M, Moss-Morris R, Caskey F, Chilcot J: Fatigue in advanced kidney disease. Kidney Int 2014;86:497-505.
2 Bossola M, Di Stasio E, Antocicco M, Panico L, Pepe G, Tazza L: Fatigue Is Associated with Increased Risk of Mortality in Patients on Chronic Hemodialysis. Nephron 2015;130:113-118.
3 Chilcot J, Moss-Morris R, Artom M, Harden L, Picariello F, Hughes H, Bates S, Macdougall IC: Psychosocial and Clinical Correlates of Fatigue in Haemodialysis Patients: the Importance of Patients' Illness Cognitions and Behaviours. Int J Behav Med 2016;23:271-281.
4 Bossola M, Vulpio C, Tazza L: Fatigue in chronic dialysis patients. Semin Dial 2011;24:550-555.
5 Picariello F, Moss-Morris R, Macdougall IC, Chilcot J: The role of psychological factors in fatigue among end-stage kidney disease patients: a critical review. Clinical Kidney Journal 2016;advanced online access
6 Chilcot J, Norton S, Kelly ME, Moss-Morris R: The Chalder Fatigue Questionnaire is a valid and reliable measure of perceived fatigue severity in multiple sclerosis. Mult Scler 2016;22:677-684.
7 Michielsen HJ, De Vries J, Van Heck GL, Van de Vijver FJR, Sijtsma K: Examination of the Dimensionality of Fatigue. European Journal of Psychological Assessment 2004;20:39-48.
8 Picariello F, Moss-Morris R, Macdougall IC, Chilcot J: Measuring fatigue in haemodialysis patients: The factor structure of the Chalder Fatigue Questionnaire (CFQ). J Psychosom Res 2016;84:81-83.
9 McCann K, Boore JR: Fatigue in persons with renal failure who require maintenance haemodialysis. J Adv Nurs 2000;32:1132-1142.
10 O'Sullivan D, McCarthy G: An exploration of the relationship between fatigue and physical functioning in patients with end stage renal disease receiving haemodialysis. J Clin Nurs 2007;16:276-284.
11 Smets EM, Garssen B, Bonke B, De Haes JC: The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. J Psychosom Res 1995;39:315-325.
12 Chandel P, Sultan A, Khan KA, Choudhary V, Parganiha A: Validation of the Hindi version of the Multidimensional Fatigue Inventory-20 (MFI-20) in Indian cancer patients. Support Care Cancer 2015;23:2957-2964.
13 Elbers RG, van Wegen EE, Verhoef J, Kwakkel G: Reliability and structural validity of the Multidimensional Fatigue Inventory (MFI) in patients with idiopathic Parkinson's disease. Parkinsonism Relat Disord 2012;18:532-536.
14 Baptista RLR, Biasoli I, Scheliga A, Soares A, Brabo E, Morais JC, Werneck GL, Spector N: Psychometric Properties of the Multidimensional Fatigue Inventory in Brazilian Hodgkin's Lymphoma Survivors. Journal of Pain and Symptom Management 2012;44:908-915.

15 Lin JM, Brimmer DJ, Maloney EM, Nyarko E, Belue R, Reeves WC: Further validation of the Multidimensional Fatigue Inventory in a US adult population sample. Popul Health Metr 2009;7:18.
16 Friedli K, Almond M, Day C, Chilcot J, Gane Mda S, Davenport A, Guirguis A, Fineberg N, Spencer B, Wellsted D, Farrington K: A study of sertraline in dialysis (ASSertID): a protocol for a pilot randomised controlled trial of drug treatment for depression in patients undergoing haemodialysis. BMC Nephrol 2015;16:172.
17 Friedli K, Guirguis A, Almond M, Day C, Chilcot J, Gane Mda S, Davenport A, Fineberg NA, Spencer B, Wellsted D, Farrington K: Sertraline versus placebo in patients with major depressive disorder and undergoing haemodialysis (ASSertID): a randomised, controlled feasibility trial. Clin J Am Soc Nephrol 2017;12:280-286.
18 Beck AT, Steer RA, Brown GK: Beck Depression Inventory—2nd Edition Manual. San Antonio, The Psychological Corporation, 1996.
19 Spitzer RL, Kroenke K, Williams JB: Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. JAMA 1999;282:1737-1744.
20 Chilcot J, Wellsted D, Farrington K: Screening for depression while patients dialyse: an evaluation. Nephrology Dialysis Transplantation 2008;23:2653-2659.
21 Chilcot J, Norton S, Wellsted D, Almond M, Davenport A, Farrington K: A confirmatory factor analysis of the Beck Depression Inventory-II in end-stage renal disease patients. J Psychosom Res 2011;71:148-153.
22 Watnick S, Wang PL, Demadura T, Ganzini L: Validation of 2 depression screening tools in dialysis patients. American Journal of Kidney Disease 2005;46:919-924.
23 Ullman JB: Structural equation modeling: reviewing the basics and moving forward. J Pers Assess 2006;87:35-50.
24 Hu L, Bentler PM: Cuttoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling 1999;6:1-55.
25 Flora DB, Curran PJ: An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. Psychol Methods 2004;9:466-491.
26 Chilcot J, Wellsted D, Da Silva-Gane M, Farrington K: Depression on dialysis. Nephron Clin Pract 2008;108:c256-264.
27 MacNeill SJ, Casula A, Shaw C, Castledine C: UK Renal Registry 18th Annual Report: Chapter 2 UK Renal Replacement Therapy Prevalence in 2014: National and Centre-specific Analyses. Nephron 2016;132(suppl 1):41-68.
28 Steenkamp R, Caskey F: UK Renal Registry 18th Annual Report: Chapter 6 Comorbidities and Current Smoking Status amongst Patients starting Renal Replacement Therapy in England, Wales and Northern Ireland from 2013 to 2014. Nephron 2016;132(suppl 1):145-154.

Table 1: Means, intercorrelations and internal reliabilities for the MFI-20 sum scores.

			MFI-20 Sum scores (correlation matrix)			
MFI-20 sum scores	Mean (s.d.)	Cronbach's alpha	**1**	**2**	**3**	**4**
1. General	13.1 (4.0)	.66	1			
2. Physical	14.5 (4.2)	.71	.670**	1		
3. Mental	9.6 (4.4)	.73	.458**	.351**	1	
4. Reduced Activity	13.1 (4.0)	.55	.983**	.647**	.445**	1
5. Reduced Motivation	13.2 (4.2)	.75	.958**	.642**	.432**	.965**

**p<0.01

S.D = Standard deviation

Cronbach's alpha = measure of internal reliability

Table 2: Summary of MFI-20 confirmatory factor models

Model	**Description**	**No of free parameters**	**Chi-square (df), p-value**	**CFI**	**TLI**	**RMSEA**
A	Original 5 factor	70	919.8 (160), p<0.01	0.738	0.689	0.101
B	1-factor	60	1181.2 (170), p<0.01	0.652	0.611	0.112
C	Bi-factor	74	912.4 (156), p<0.01	0.739	0.683	0.102
D	3-factor	63	985.0 (167), p<0.01	0.718	0.679	0.102
Original 5-factors[a]						
General		12	25.6 (2), p<0.01	0.892	0.677	0.159
Physical		12	16.2 (2), p<0.01	0.941	0.822	0.123
Mental		12	33.5 (2), p<0.01	0.889	0.668	0.182
Reduced activity		12	23.2 (2), p<0.01	0.930	0.789	0.150
Reduced motivation		12	73.4 (2), p<0.01	0.485	0.546	0.276

Root Mean Squared Error of Approximation (RMSEA); Confirmatory Fit Index (CFI); Tucker-Lewis index (TLI).

Degrees of Freedom (df).

[a]Each factor of the original MFI-20 was analyzed in independent models.

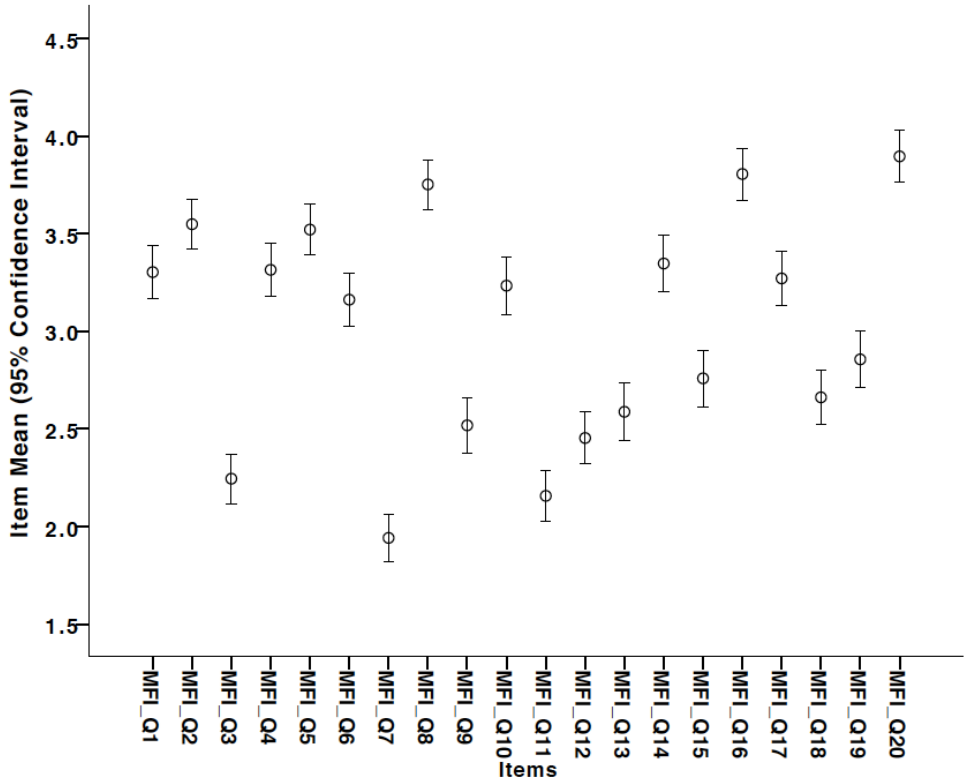

Figure 1: MFI-20 individual item score means with 95% confidence intervals for the item means